



Grid Analysis of Radiological Data

Cecile Germain-Renaud, Vincent Breton, Patrick Clarysse, Bertrand Delhay, Yann Gaudeau, Tristan Glatard, Emmanuel Jeannot, Yannick Legre, Johan Montagnat, Jean-Marie Moureaux, et al.

► To cite this version:

Cecile Germain-Renaud, Vincent Breton, Patrick Clarysse, Bertrand Delhay, Yann Gaudeau, et al.. Grid Analysis of Radiological Data. Mario Cannataro (Ed.). Handbook of Research on Computational Grid Technologies for Life Sciences, Biomedicine and Healthcare, IGI, pp.363-391, 2009, chapter 19, 10.4018/978-1-60566-374-6.ch019 . hal-00683992

HAL Id: hal-00683992

<https://hal.science/hal-00683992>

Submitted on 30 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Grid Analysis of Radiological Data

Cécile Germain-Renaud

*Laboratoire de Recherche en Informatique (Univ. Paris-Sud; CNRS; INRIA)
France*

Vincent Breton

*Laboratoire de Physique Corpusculaire de Clermont-Ferrand (CNRS; Univ. Blaise Pascal)
France*

Patrick Clarysse

*CREATIS (INSA Lyon; CNRS; Univ. Claude Bernard Lyon 1; INSERM)
France*

B. Delhay

*CREATIS (INSA Lyon; CNRS; Univ. Claude Bernard Lyon 1; INSERM)
France*

Yann Gaudreau

*Centre de Recherche en Automatique de Nancy (Univ. Henri Poincaré; CNRS)
France*

Tristan Glatard

*Laboratoire d'Informatique, Signaux et Systèmes (CNRS; Univ. Nice Sophia-Antipolis; INRIA)
France*

Emmanuel Jeannot

*Laboratoire Lorrain de Recherche en Informatique et ses Applications (INRIA; CNRS; Univ. Henri Poincaré)
France*

Yannick Legré

*Laboratoire de Physique Corpusculaire de Clermont-Ferrand (CNRS; Univ. Blaise Pascal)
France*

Johan Montagnat

*Laboratoire d'Informatique, Signaux et Systèmes (CNRS; Univ. Nice Sophia-Antipolis; INRIA)
France*

Jean Marie Moureaux,

*Centre de Recherche en Automatique de Nancy (CNRS; Univ. Henri Poincaré)
France*

Angel Osorio

*Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (CNRS)
France*

Xavier Pennec

*INRIA Sophia-Antipolis
France*

Joël Schaerer

*CREATIS (INSA Lyon, CNRS, Univ. Claude Bernard Lyon 1 INSERM)
France*

Romain Texier

*Laboratoire d'Informatique, Signaux et Systèmes (CNRS ; Univ. Nice Sophia-Antipolis;
INRIA)
France*

Corresponding author: cecile.germain@lri.fr

Abstract

Grid technologies and infrastructures can contribute to harnessing the full power of computer-aided image analysis into clinical research and practice. Given the volume of data, the sensitivity of medical information, and the joint complexity of medical datasets and computations expected in clinical practice, the challenge is to fill the gap between the grid middleware and the requirements of clinical applications. This chapter reports on the goals, achievements and lessons learned from the AGIR (Grid Analysis of Radiological Data) project. AGIR addresses this challenge through a combined approach. On one hand, leveraging the grid middleware through core grid medical services (data management, responsiveness, compression, workflows) targets the requirements of medical data processing applications. On the other hand, grid-enabling a panel of applications ranging from algorithmic research to clinical use cases both exploits and drives the development of the services.

INTRODUCTION

Harnessing the full power of computer-aided image analysis into clinical research and practice remains an open issue. Given the amount of data produced by X-ray Computed Tomography (CT), Magnetic Resonance Imaging (MRI), or PET-scan, and the difficulty to interpret medical images, algorithms for medical image analysis, processing, and diagnostic assistance have been developed these last 15 years or so. Some of these algorithms have reached a high level of usability and proved to have a real impact in the clinical domain. However, their widespread adoption by clinicians is not realized yet. Two stringent examples, amongst many others, are radiotherapy, which could greatly benefit of exploiting advances in segmentation and registration algorithms, and intra-operative situations as well as intervention planning, which could exploit modern high-performance computing systems for augmented reality (Kikinis, 1998). In the 90's, G.A Moore has coined the term "crossing the chasm" (Moore, 1991) for the issue of selling high-tech products to mainstream customers. Crossing the long-lasting chasm between, on one hand the advances in computer science and engineering in the field of medical images analysis, and on the other hand clinical research and practice, is a challenge of the same nature.

This chapter reports on the goals, achievements and lessons learned from the AGIR (Grid Analysis of Radiological Data) project. AGIR is a multi-disciplinary collaboration funded by the French ministry of research under the ACI/ANR scheme, for the 2004-2007 period. AGIR gathers

collaborators from eight laboratories, including computer scientists, middleware experts and physicians. The central tenet of AGIR is that grids can help crossing the chasm, despite their recent apparition, because they introduce a change in paradigm in the access to high-end resources. Grids provide seamless scaling of the development, validation, and exploitation cycles of analysis methods: the same infrastructure allows multiple user communities to access and manipulate medical data, and to explore individual images or create augmented reality situations through compute-intensive visualization methods; the same infrastructure provides the computing power needed to validate algorithms on large datasets and to process complete databases for the most demanding applications such as epidemiology.

Filling the gap between the clinical applications and the grid middleware raises many specific issues, ranging from computer science basic research to legal concerns, as will be exemplified in section 2. Addressing all these issues is an active research and technology area, and a new scientific community is emerging (Berry, 2003). In this framework, the specific objectives of AGIR are to define and validate:

- new grid services that address some of the requirements of complex medical image processing and data manipulation applications; these services are described individually in section 3.
- new medical image processing algorithms taking advantage of the underlying grid infrastructure for compute and data intensive needs; sections 4 to 6 report on these developments.

The method is to confront the expertise of medical and computer science teams, specialized in clinical applications, medical images analysis algorithms, grids and distributed systems, around a few paradigmatic medical applications, in order to get a cross-section of the middleware, algorithmic and medical issues. A specific application (section 7) was a “bottom-up” approach targeting the immediate development of a telemedicine platform, whose requirements and results were to be confronted with the more long-term activities in AGIR.

GRID-ENABLING MEDICAL IMAGE ANALYSIS

Challenges and issues

Grid computing has been considered to tackle some of the requirements of the medical image analysis community (Montagnat, 2004; Aloiso, 2005). The scenarios investigated in AGIR (see section 4-6) led to the identification of common area-specific grid services:

- Medical databases management in order to handle large amounts of sensitive data. Federation of distributed data sources is needed e.g. for large scale testing involved in algorithms validation.
- Optimized and reliable data transfers even on low-end networks, e.g. to support humanitarian medicine.
- Efficient execution of analysis pipelines on a distributed grid infrastructure in order to match the computing requirements of the most demanding applications such as cardiac sequences analysis.
- Responsiveness needed by human guided procedures such as assisted segmentation and surgery planning.
- Production quality support to ensure sustainable grid usability.

Overview of AGIR

The deployment of medical applications on grids can be structured following four layers: 1) general purpose grid services; 2) the core services dedicated to medical applications that are not

available in general purpose middleware; 3) the medical image processing algorithms taking advantage of the underlying grid services to process large amounts of data or to run expensive computations; 4) clinical applications developed to address medical challenges.

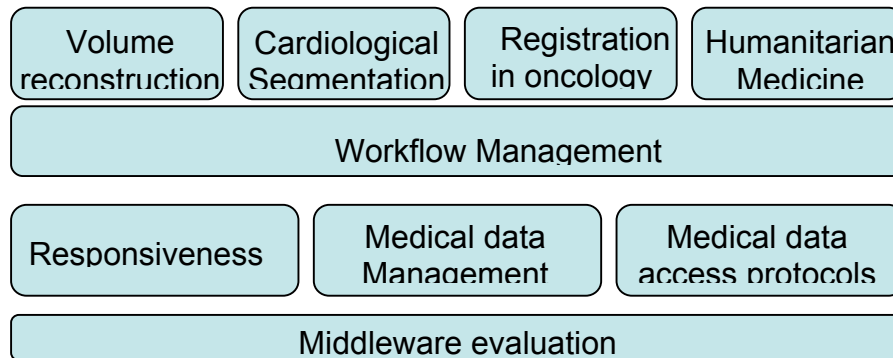


Figure 1: AGIR structure

AGIR mainly addresses core medical services and image processing algorithms and, to some extent, clinical applications. Figure 1 summarizes the various AGIR activities.

e- infrastructures

Grid concepts and research are around since the 90's at least (Foster, 2003) (Foster, 2001). With the maturation of technologies, and the deployment of production grids, however, the focus has now shifted towards strategic issues, amongst which forecasting and shaping roadmaps towards future e-infrastructures (Atkins, 2003). Following the analysis of the e-IRG (e-infrastructures reflection group, a body of EU member states representatives advising the European Commission), sustaining the needs of science requires resource integration at the European scale not as a fully integrated ("monolithic") grid, but at the level of an interoperable grid of grids (Leenars, 2005).

At the beginning of AGIR, these issues were not fully formulated. The popularity of the Globus middleware (Foster, 2006) pushed towards autonomous developments. However, we already considered that it was critical for AGIR to refrain from developing middleware functionalities, which were already available, even if not perfectly suited for its needs. Throughout the project, we tried to keep a balance between neutrality with respect to middleware implementations on one hand, and strategic concerns on the other hand.

On the generic side, we conducted experiments and analysis in order to assess the compatibility of various middleware platforms, from research ones to production oriented ones. Amongst the numerous research suites, our experiments focused on DIET (The Distributed Interactive Engineering Toolbox) (Caron, 2006), which exemplifies remote procedure call (RPC)-based architectures. Another important component of our experiments is the French national grid, Grid'5000 (Bolze, 2006), which allows repeatable experiments in controlled settings.

On the strategic concerns side, AGIR has developed a close collaboration with the FP6 and FP7 infrastructure projects EGEE (Gagliardi, 2005), which is expected to move towards a permanent European Grid Infrastructure federating National Grid Infrastructures (NGI). EGEE is strongly involved in the various prospective bodies such as e-IRG, interoperability actions with other international grids (e.g. NorduGrid or Open Science Grid), and with the standardization

efforts of the Open Grid Forum. EGEE is thus one of the key actors in the European e-infrastructures theater.

EGEE is a controlled but widely distributed environment over the Internet which usage for manipulating sensitive medical data imply the implementation of strong protection measures. However, the strong interaction with EGEE has a triple advantage. The first and most immediate one is the sheer availability of resources, showing that IT resources are readily available, which is a cornerstone for real-world usage scenarios. The second advantage is to interact with middleware developments, specifically gLite (Laure, 2004). On one hand, this process is considerably more difficult than it would be in an experimental grid. On the other hand, the impact of these requirements, when satisfied, benefits from the general impact of EGEE, and are sustainable.

An alternative might be to separate grids for health from other e-infrastructures. In the AGIR time frame, other large scale infrastructures targeting life sciences have been deployed world-wide, adopting this strategy. In particular, the BIRN (Biomedical Informatics Research Network) project (Peltier, 2003) has deployed an homogeneous hardware and software infrastructure all over the USA. Similarly on the EU side, Health-e-child¹ has deployed an EGEE-compatible but independent infrastructure. This has the advantage to minimize the legal and ethical issues as the nodes are owned and controlled by the hospitals themselves so that they comply to each site requirements. Moreover, the different sites are much more easily administrated by the consortium. We believe that this is a good way to expose the clinical community to grid technologies in order to facilitate in the future the acceptance and the use of more standard e-infrastructures. Conversely, the CaBIG² project has adopted a more open approach with the publication of different level compatibility recommendations for the different components, existing or to appear, of its grid infrastructure.

In both cases the logic of a production grid is the same: achieving a proper balance between, one hand quality and sustainability (24x7 availability, and compliance to standards), and on the other hand coping with scaling requirements, both in heterogeneity (workloads and use cases), and in geographical and institutional distribution.

CORE MEDICAL SERVICES

Grid interface for Medical Data Management

Medical data management in clinics is a complex problem given the multiplicity of data sources, the sensitivity of the data, and the tremendously large data sets involved. As a consequence, medical information systems are often fragmented between different clinical sites where homogenization efforts are rarely invested. Even inside a single health enterprise, it is not rare to encounter several information systems partially overlapping and covering the different aspects of clinical information with more or less success.

Considering medical images, we have developed a complete solution, the *Medical Data Manager* (MDM). The MDM, interfaced to the EGEE middleware, bridges the DICOM storage used internally to clinical centres and the grid storage units (Montagnat, 2008). The MDM both ensures a transparent access to medical images by any grid service and does not interfere with the usual clinical workflow of images. Benefiting from the distributed storage capability provided by the grid middleware, it enables data sources federation: large image sets originating from different sites can be assembled for studies involving processing population of images such as the algorithm validation through statistical procedures as the Bronze Standard (section 5).

¹ Health-e-child: Integrated platform for paediatrics, <http://www.health-e-child.org/>

² CaBIG: Cancer Bioinformatics Grid, <https://cabig.nci.nih.gov/>

Contexts

In health enterprises, the Digital Image and COmmunication in Medicine (DICOM) standard has become by far the widely available interface to access medical storage. Although a largely adopted standard, DICOM poorly matches the need for cross-health enterprise image exchanges. It was seminally developed for local network exchanges. It is known to provide both a low level of access control and a low efficiency in data transfers. Grids technologies can help on this issue: they not only provide federated data services but also reliable, secure and efficient protocols to transport data over the open Internet. However, a issue specific to medical data management is the fact that privacy preserving at the grid level must not hamper possible patient benefit. The challenge is thus to harness the DICOM access with the grid capacities regarding security, performance and storage resources.

DICOM is exploited by image device acquisition and post-processing vendors to ensure image communications in the clinical Picture Archiving and Communication Systems (PACS). DICOM defines both an image format and a transfer protocol which are completely integrated in standard clinical image workflows. DICOM stores structure medical images hierarchically in studies, series and individual image slices. On the other hand, grid technologies provide distributed file catalogs and remote file storage access protocols suited to build a file-based distributed data management system.

The Medical Data Manager

The MDM was designed to interface the EGEE grid infrastructure with state of the art DICOM compliant storage resources inside hospitals. It mostly targets the exploitation of medical image files as manipulated by image processing algorithms. It exploits a limited amount of associated metadata though to ease image retrieval in clinical use cases.

All medical data is considered as sensitive to preserve patient privacy. Nominative medical records are of course the most critical data and therefore, no binding between nominative data and images should be possible for non accredited users. On a grid, data distribution aggravates the security problems. To ensure patient privacy, the header of all DICOM images sent out of a DICOM server is removed, at least partially, to ensure patients' anonymity. All images stored out of the source center is encrypted to ensure that non accredited users cannot read the image content.

To fulfill both the needs of clinical practice and grid computing, the MDM:

- Provides access to medical data sources for computing without interfering with the clinical image workflow.
- Ensures transparency so that accessing medical data does not require any specific user intervention.
- Ensures a high data protection level to respect patients' privacy and data confidentiality constraints.

The MDM is thus compliant with DICOM clinical servers internally and interfaced to the EGEE grid data management system externally.

gLite data management layer

The gLite middleware deployed on EGEE provides data management services for handling distributed files. Storage resources, known as *Storage Elements*, are distributed over the various sites participating in the infrastructure. All storage elements expose a common interface for interacting with the other middleware services: the *Storage Resource Manager* interface (SRM) that is standardized in the context of the Open Grid Forum (Sim, 2007). The SRM is handling local data at the file level. It does not implement data transfer by itself: additional services such as

GridFTP (Allcock, 2002) or gLiteIO (EGEE JRA1 2005b) are coexisting within storage elements to provide transfer and data access capabilities. The gLite data management system uses a *File Catalog* providing a unique entry point for files distributed on all grid storage elements. Each file is uniquely identified through a *Global Unique Identifier* (GUID). The file catalog contains an association between GUIDs and site-dependent Storage URLs (SURL) enabling distributed files location. The File Catalog and the Storage Element may implement different file access policies and it is one of the MDM challenge to ensure a coherent grid-wide access control through coordinated updates of the fine-grain Access Control Lists (ACLs) handled by the different services.

MDM services

The MDM is a data management layer based on multiple grid services. It uses:

- A core MDM component exposing on SRM interface and querying the DICOM server.
- An ACL-based access right controller.
- A file catalog to index images available in the DICOM server.
- A secured file transfer service for sending images to requesting grid services.
- An access-controlled key store associating encryption keys to images.

In addition, the AMGA secured relational database interface is used for managing metadata associated to images (such as image metadata, patient metadata, radiology metadata, etc). AMGA (Santos, 2006) is a secured front-end to traditional databases back-ends (MySQL, Postgres, Oracle). It provides grid credential ACL-based access control and a secured communication layer. It relates medical information with grid file identifier (GUID).

MDM security

The latest version of the MDM is based on the lightweight *Disk Pool Manager* (DPM) disk server which implements SRM version 2 (Perelmutov 2002). Access is controlled through the combined use of the *LFC* File Catalog and the back-end SRM server. The gLite secured transport protocols (typically GridFTP) are used for data transfer. Encryption keys are stored in a *Hydra* distributed keystore exploiting the Shamir secret sharing scheme (Shamir, 1979). Shamir's scheme consists in splitting keys into n fragments stored in different places. Only $m < n$ fragments are needed to reconstruct a complete key. Thus, the system security is improved as it is both resistant to attacks and provides redundancy. The standard symmetric key algorithm using AES cipher available in the OpenSSL distribution is used for encryption/decryption. All images are encrypted prior to communication thus ensuring on-transmission security and on-disk security once the data file is stored on a remote site. Only users with access right to the file key can access the data content. In addition, DICOM headers which contain nominative information are anonymized prior to transfer.

The MDM provides a much higher level of data security than the standard gLite Data Management System. It required the addition of specific components such as the coordinated access control and the Hydra key store for encryption.

MDM service set up and exploitation

To fulfill its role, the MDM service needs to be notified when files are produced by the imagers and stored into the DICOM server. The DICOM data triggering the operation is first stored into the hospital DICOM server as usual. The DICOM header is then analyzed to extract image identifying information. This DICOM ID is used to build a file identifier that is registered into the grid File Catalog. The other metadata extracted from the DICOM header is stored into the AMGA metadata server. Finally, an encryption key that is associated to the file and that will be used for data retrieval is stored into the Hydra key store.

Once DICOM data sets have been registered into the MDM, the server is able to deliver requested data to the grid. A client library is used for this purpose. To cover all application use cases, the MDM client library provides APIs for requesting files based on their grid file identifier or on the metadata attached to the file. In many cases, complete image volumes are needed rather than individual image slices. A grid file identifier is associated to each volume registered and the image volume can be reconstructed on the fly by the SRM-DICOM core on request. All image files are anonymized and encrypted before being sent out to ensure that no sensitive information is ever transferred or stored on the grid in a readable format.

Discussion

The MDM service is a gLite integrated DICOM-compliant storage resources manager. It provides transparent access to the clinical storage from other gLite services and it ensures a high security level. A similar service has been recently developed on the GLOBUS middleware stack. The MEDICUS incubator project (Erberich 2007a) is proposing a DICOM interface to the GLOBUS data manager. MEDICUS is based on OGSA-DAI compliant Meta Catalog Service (MCS) and a DICOM Grid Interface Service (DGIS). As compared to the MDM, MEDICUS does not directly index DICOM storage images into the file catalog. Instead, it explicitly exports images selected for grid access into the DGIS server deployed in demilitarized zone. The data is therefore replicated on a specific grid storage resource before being accessible. MEDICUS does not provide file encryption facility either. It relies on the GridFTP transport layer to ensure file security upon transfer. It does not provide on-disk file encryption though. MEDICUS provides an elaborated file access control mechanism through the use of Shibboleth (Erberich 2007b).

Responsiveness

Responsiveness is the support for on-demand access and interactive usage (Basu, 2003). The choice of targeting production grids makes responsiveness most challenging. Interaction and on-demand features must be built on top of the traditional grid scheduling tools, which are batch-oriented and dominated by fair-share policies at institutional time-scales. The associated constraints might be summarized as non-intrusiveness: delays incurred by non-interactive jobs must be bounded, resource utilization should not be degraded (e.g. by idling processors), and finally, the local policies governing resource sharing (Virtual Organizations, advance reservation, etc.) should not be impacted. Another related issue is the middleware latency: even if all the necessary resources are available, submitting, scheduling and mapping of jobs on EGEE remains in the order of minutes. Thus, two orthogonal issues must be addressed.

- Differentiated Quality of Service (QoS) at the workload management level, where batch jobs are serviced as best effort, and responsive applications are guaranteed to be serviced without queuing .
- Coping with the middleware penalty, which requires to bypass the middleware stack.

Differentiated QoS

Intra-operative situations require advance reservation. For more routine clinical or research grid usage, advance reservation suffers from two drawbacks: first, planning is not consistent with the goal of seamless integration with everyday computing practice, and specifically medical usage; second, reservation is inherently not work-conserving, meaning that processors might idle while eligible jobs are queued (Snell, 2000), thus jeopardizing utilization. Differentiated QoS is the generic answer, and has been extensively explored in both the network and operating systems (OS) frameworks. Differentiated QoS usually relies on some implementation of Generalized Processor Sharing (GPS)(Baruah, 1996), where resources may be allocated by quanta. In grids, such quanta do not exist: except for the rare case of checkpointable jobs, a job that has started

running cannot be suspended and restarted later; moreover, the execution times exhibit high variance (Germain, 2006).

We have defined and implemented the concept of *Virtual Reservation* (VRes) (Germain, 2006), which addresses both issues of advance reservation and scheduling quanta by allowing controlled time-sharing. VRes permits the definition of time quanta and their exposure at the grid level. At the site level, each of the p physical processors is virtualized into k virtual processors, providing pk slots to the site scheduler. When a virtual slot is unused, the computing bandwidth is transparently returned to the other classes sharing the same physical processor, by the virtue of time-sharing. Thus, a fraction of these slots can then be permanently reserved for the interactive applications without jeopardizing utilization. One of the advantages of this scheme, is that classes (batch and interactive) are isolated: resource allocation becomes internal to the classes. True QoS has the unavoidable counterpart of access control: the guaranteed bandwidth being limited by the number of available slots, a job requiring QoS may be rejected if the available resource is exhausted.

An implementation of VRes has been developed for the MAUI scheduler and the gLite middleware (EGEE SDJ WG, 2006). The Job Description language (JDL) has been modified to include a Boolean attribute SDJ (Short Deadline Job), denoting that the job must be executed immediately, or rejected. These sites also configure their scheduler with parameters controlling the computational bandwidth dedicated to SDJ. VRes has been put in production since May 2006. and is routinely used by several biomedical applications.

User-level scheduling

Submitting, scheduling and mapping of jobs on a grid take at least one order of magnitude more time than the execution time for SDJ even in absence of competition for resources. User-level scheduling is the most promising way to address this issue. User-level scheduling is a virtualization layer: instead of being executed directly, the application is executed via an overlay scheduling layer, which runs as a set of regular user jobs and therefore bypasses the most costly middleware services. User-level is here analogous to the user-level thread schedulers versus kernel-level ones in processor scheduling. Just as a user-level thread scheduler avoids the penalties of the kernel calls, a Grid application-level scheduler avoids the traversal of the Grid information systems for executing each and every small parallel task.

AGIR has embedded a user-level scheduler inside the gPTM3D application, corresponding to the interactive volume reconstruction use case. User-level scheduling can instead be factorized across applications through generic frameworks, also called overlay grids, such as AppLeS (Berman, 2003), DIRAC (Tsaregorodtsev, 2004), DIANE, or the more complete solution DIET presented above. While DIRAC is not integrated with gLite, which precluded its usage in AGIR, DIANE is. It would have thus be useful to capitalize over the DIANE generic framework. However, synthetic experiments (Texier, 2007) have shown that the latency of DIANE remains too high for the volume reconstruction use case. Finally, we have been able to experiment the DIET scheduler. A detailed experiment (Texier, 2007) run on the Grid 5000 resource has proved that DIET provides low latency for fine-grained tasks.

Discussion

An alternative approaches to grid scheduling for QoS is virtualization; the Virtuoso project has more specifically explored virtualization for differentiated services (Lin, 2005). Virtual machine can indeed be deployed in stable grid environments, e.g. the Virtual Workspaces project (Keahey, 2006) with Globus. Virtual machines provide complete freedom of scheduling and even migrating an entire OS and associated computations which considerably eases time-sharing

between deadline-bound short jobs and long running batch jobs. On the other hand, the virtual machines strategy requires that all of, or a significant fraction of, the computations must be run inside virtual machines to provide scheduling opportunities, something for which traditional batch users have little incentive. As the role of interactive users is still modest in the grid community, the smoother path proposed by VRes represents a more acceptable tradeoff.

Compression

Transmission as well as storage of the tremendous amount of data involved in medical applications can easily saturate existing systems. Along the interaction of AGIR with clinicians, it appeared that a prominent requirement for any networked access to images is low response time. In the long-term vision of grids as the infrastructure for perennial conservation and pervasive access to medical data, compression is thus an unavoidable component of medical data access protocols.

In the following we give details about the two components of enhanced medical data access protocols we have explored: lossy compression algorithms, which are required for efficiency, but yet questioned by the medical community, and adaptive coupling of compression and transfer.

Compression algorithms

The tolerance to lossy compression (defined as the maximal threshold of compression for which the decompressed image is still acceptable for interpretation) is related to the type of data. Digitized chest radiographs are very tolerant to compression (Erickson, 2002), at least 40:1 for 2D SPIHT (Cho, 2004), digitized bone films are moderately tolerant (between 20:1 and 40:1), and CT or MR images exhibit fairly low tolerance to compression (less than 20:1) (Erickson, 2002). In the following, we consider 3D images, or more exactly stacks of 2D images, like MRIs or CTs.

The basic idea of 3D medical image coding algorithms is to take advantage of the correlation among the data samples in the three dimensional space to improve compression performance. Most of the approaches combine a three-dimensional space decorrelating transform with the extension of a coding algorithm that has proven to be effective on 2D images. The reference methods are based on 3D Wavelet Transform (Bilgin, 2000).

Two strategies are available to encode wavelet subbands supplied by the 3D WT: the interband and the intraband approaches respectively. While the interband schemes, exemplified by 3D SPIHT (Cho, 2004), are generally more flexible and offer a lower computational cost than intraband schemes, they are sensitive to errors occurring during data transmission.

On the contrary, the intraband schemes encode the whole subbands separately and need a bit allocation procedure to share the binary resources among the subbands in order to minimize the overall distortion. Recently, we have proposed a new multidimensional intraband algorithm based on lattice vector quantization associated with a vector dead zone: 3D DZLVQ (Gauzeau, 2007). The corresponding codebook structure take into account both the sparsity and the clustering properties which characterize the wavelet coefficients. Thanks to the proposed dead zone, this scheme allows to threshold non significant source vectors according to a criterion based on their norm, while putting more bits on significant ones. It can efficiently apply to medical images in the 3D wavelet domain which often contain huge non-significant areas.

	3D SPIHT	3D DZLVQ	3D ESCOT
0.5 bpv	44.11 dB	44.14 dB	43,82 dB
0.1bpv	33.98 dB	34.69 dB	34,68 dB

Table 1. PSNR (dB) comparison on slice 128 of Skull image compressed by 3D SPIHT, 3D DZLVQ and 3D ESCOT at two bit rates: 0.5 bit/voxel and 0.1 bit/voxel.

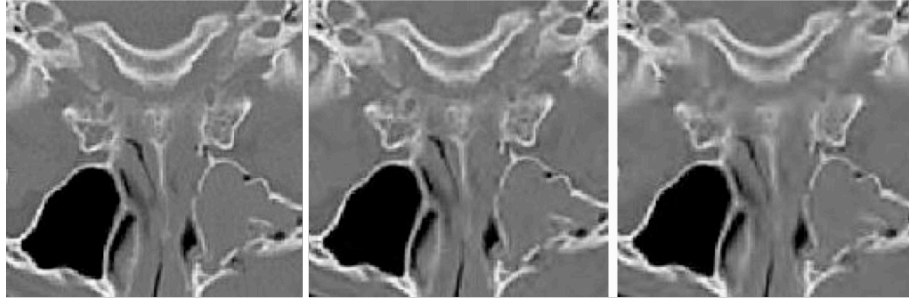


Figure 2. Zoom of decoded slice 60 (Skull) when the target decoding bit rate is 0.1 bpv for the whole 128-slice unit. From left to right: Original slice, 3D SPIHT, 3D DZLVQ.

Table 1 shows compression performance in terms of Peak Signal-to-Noise Ratio (PSNR) of DZLVQ compared to 3D SPIHT, as well as 3D ESCOT (Xiong, 2003), one of the most efficient intraband algorithms of the literature. The selected 256x256x192 8-bit CT image “Skull” of this experiment belongs to an abundantly used volumetric medical image data set given in (Bilgin, 2000). Furthermore, a four level 3D WT has been performed, using the well-known 9.7 floating-point filter. As we can see in the table, 3D DZLVQ slightly outperforms 3D SPIHT and 3D ESCOT. Note that the same tendency has been observed for the other images of the data set and for DICOM 12-bit ENT-CTs (Gaudeau, 2007).

PSNR, as a global criterion, is not sufficient to estimate the quality of reconstructed images, especially in medical applications. Figure 2 shows that 3D DZLVQ allows to better preserve fine structures and produces a better overall visual quality of the reconstructed image (Gaudeau, 2007).

Adaptive On-line Compression of Medical Images

While local area networks (e.g. hospital intranets) offer the bandwidth required for quasi instantaneous transfer of 3D images, widening the scope of access may involve poor (e.g. roaming user) or congested (e.g. wide area networks) infrastructures, resulting in low bandwidth. In a previous work, we developed AdOC (Adaptive Online Compression) (Jeannot, 2005), which enables on-the-fly compression, thus reducing the required bandwidth.

In AGIR, we have combined the principles of AdOC and wavelet transform based lossy compression. In this framework, each wavelet coefficient of a generated subband is approximated by quantization and then encoded. The adaptive techniques of AdOC are exploited to change the coding effort depending of the environment. If the network is very slow, more aggressive coding algorithms are used to compress the coefficients. Therefore, we are able to detect the best tradeoffs between spending some time for compressing the data and sending them to the network as fast as possible. Furthermore, in order to speed-up the whole process of compression and decompression we have pipelined each step (wavelet transform, quantization and coding); on a multi-core architecture, which is now the standard processor, these different steps are overlapped and data start to be sent to the network only after a minimum processing of the image.

We have enhanced the proposed framework to the case of DICOM 3D images. Each image of the 3D stack is processed separately in a pipelined manner. We have found that doing this way is much faster than performing 3D wavelet transform and enables real-time transfer.

We have tested this implementation on several images and several kinds of networks. Preliminary results show that the proposed approach is very promising for wide-area networks such as Internet both in terms of transfer speed (it can be divided by 3 on transatlantic connection compared to FTP transfer) and quality the images (they do not suffer visible degradation). Moreover, the progressive transmission of images allows the clinician to discard those of less interest without a complete upload.

Image pipelines and parallel processing

A pipeline, better known as workflow in the grid community, enables the definition of processing chains composed of different image processing units (e.g. image filtering, intensity normalization, registration, segmentation, parameters quantification, interpretation...) and dependencies between these units. The dependencies constrain the order of execution of the pipeline's units. In the context of grid computing, these pipelines are also a convenient and user friendly way to decompose a complex application into more or less independent computing units which execution can be distributed over different computing resources.

In the context of AGIR we adopted a data-flow approach to describe workflow constraints that matches the traditional notion of image analysis and visualization pipelines well known from the medical image analysis community. In a data-flow, constraints between processing units are expressed through data links describing how the output of one processing unit are transferred to the input of a subsequent processing unit. Those data transfers enforce a partial order of execution for all units. Another critical requirement for the AGIR workflow manager was the ability to embed image processing units from various origins into the same application. We adopted a flexible service-oriented architecture to match this requirement. Business codes are embedded into a Web-Service (WS) wrapper and the AGIR workflow manager invokes the WS composing the pipeline in turn.

Workflow representation

To describe AGIR workflows, we adopted Scufi (Oinn 2004), a data-flow language that combines a rich semantics a compact and user friendly language. Scufi features data composition operators, also called iteration strategies. Two main composition patterns are *one-to-one* composition pattern (left of figure 3), where two input data sets are processed pairwise in their order of arrival and the *all-to-all* composition pattern (right of figure 3), where all inputs in one data set are processed with all inputs in the other one. A common example is the case where all images in the first input set are to be processed with all parameter configurations defined in the second input set.

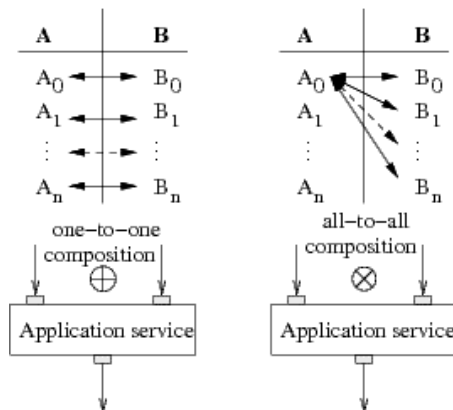


Figure 3. Action of the one-to-one (left) and all-to-all (right) operators on the input data sets.

Enacting Workflows

The Scufl workflow enactor is in charge of invoking the various services composing a workflow and ordering the transfers between consecutive services. One limitation of Scufl, bound to the nature of WS, is that each invocation targets a specific service endpoint hosted on a particular host. If the computation is data intensive, e.g. if a single service has to be invoked several times for processing different data segments, the basic WS invocation procedure overloads the host endpoint. This limitation made the execution of the data intensive medical image analysis applications far too inefficient. We developed a solution based on two components:

- GASW: a grid compliant generic web service wrapper (Glatard, 2008a).
- The MOTEUR parallel enactor for Scufl workflows (Glatard, 2008b).

GASW, a generic application code wrapper compliant with the Web-Services specification enables the execution of a legacy executable through a standard service interface. This service is generic in the sense that it is unique and does not depend on the executable code. It exposes a standard interface that can be used by any Web-Service compliant client to invoke the execution. It completely hides the grid infrastructure from the end user as it takes care of the interaction with the grid middleware. To accommodate to any executable, the generic service is taking two different inputs: a descriptor of the legacy executable command line format, and the input parameters and data of this executable. The production of the legacy code descriptor is the only extra work required from the application developer. It is a simple XML file which describes the legacy executable location, command line parameters, input and output data (including access protocols).

MOTEUR is an efficient grid execution of service workflows exploiting the parallelism levels that are explicitly or implicitly expressed in a Scufl workflow. To enable parallelism during the workflow execution, multiple application services have to be invoked concurrently. The calls made from the workflow enactor to these services need to be non-blocking in order to exploit the potential parallelism. Web Services theoretically enable asynchronous calls. However, the vast majority of existing web service implementations does not cover the whole standard and none of the major implementations provides any asynchronous service calls for now. Alternatively, MOTEUR implements asynchronous calls to Web Services at the workflow enactor level, by spawning independent system threads for each service being executed. With asynchronous calls, three levels of parallelism can be exploited:

- **Workflow parallelism** is the intrinsic parallelism expressed through the workflow graph topology. If two services lay in two independent branches of the workflow graph (i.e. there are no data links imposing an execution order constraint between these services nor their ancestors), then they can be enacted concurrently.
- **Data parallelism.** The processing of several input data segments is needed in many applications. Services can be instantiated as several computing tasks running on different hardware resources and processing different input data segments in parallel. Data parallelism is the major source of parallel execution gain in medical image analysis applications.
- **Service parallelism** corresponds to the concurrent execution of two independent data segments by two different services that are sequentially linked. This pipelining model, very successfully exploited inside CPUs, can be adapted to sequential parts of service workflows.

Fully parallel workflow execution, handling the data composition operators described above is not straightforward as the order of data items may be completely disturbed by the concurrent

execution of different jobs and a dedicated algorithm has to be designed in order to keep track of their provenance during the execution.

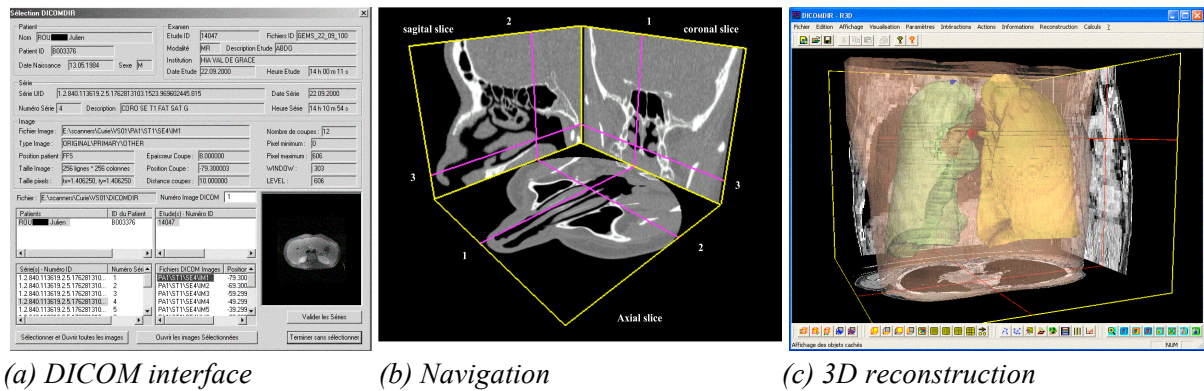
MOTEUR (Glatard 2007) is interfaced to the EGEE grid and can exploit the production infrastructure resources for running applications. The Scufi language and the MOTEUR workflow manager have been used to describe and efficiently enact various medical image analysis applications on EGEE, in particular, the rigid registration algorithms evaluation application described in section 5 (figure 6 shows the workflow).

INTERACTIVE VOLUME RECONSTRUCTION

This section exemplifies the case of grid-enabling the compute-intensive tasks involved in clinically oriented augmented reality applications.

Motivation

The original software, PTM3D, is a fully featured DICOM images analyzer developed at LIMSI (figure 4).



(a) DICOM interface

(b) Navigation

(c) 3D reconstruction

Figure 4. PTM3D snapshots.

The system currently runs on standard PC computers and it is used on line in radiology centres (Osorio, 2002). The most compute-intensive task in PTM3D is the computer-aided generation of three-dimensional representations from CT, MRI, PET-scan, or echography 3D data.

3D-reconstructions targeting surgery planning and guided surgery through augmented reality can exceed the computing power of a workstation. An example is the augmented reality system developed with PTM3D for percutaneous nephrolithotomy (figure 5) : renal stones, kidneys and external skin of the body are reconstructed from CT images; the real-time projection of organs and lesions on the patient in the surgical position leads the surgeon towards the target region and choosing the best route for percutaneous surgery (Traxer, 2005).



Figure 5. Augmented reality in surgical planning for percutaneous nephrolithotomy.

This kind of application does not require a grid as absolutely as e.g. analysis of distributed databases of medical images, because data are not initially distributed, and the computation is a classical parallel one. The motivation for seeking for grid resources is thus not different from many other medical or non-medical fields: to propose a high-end services that can be accessed freely by any authorized organization, and which provides the appropriate QoS without the burden of maintaining the resource at the individual organization level.

Requirements

The volume reconstruction algorithm includes a semi-automatic segmentation component based on an active contours method where the user initiates the segmentation, and can correct it anytime. It also includes a tessellation component, which is the compute-intensive part of the algorithm, and requires fine-grained parallelism.: the execution time of the majority of the tasks is in the order of a few hundreds of milliseconds but with high variability: when the geometry of the volume becomes complex, the reconstruction of the critical slices can last for 20 seconds or more.

The requirements are quite new with respect to the grid most frequent use cases. The challenge is to smoothly integrate local and remote computing. The radiologist wants a unique tool, to browse the images, to measure a tumor, to reconstruct a complex organ or to create a virtual reality situation. The grid must be only an accelerator of the tool.

Method

A data-driven application level scheduler has allowed fully interactive remote execution of the volume reconstruction module on the EGEE grid. This development shows that complete grid accountability and compliance with the gLite middleware is compatible with keeping sophisticated and medically-oriented interaction schemes unmodified. As we explained in the section 3.2, the generic user-level scheduler DIET offers satisfactory performance. However, the most difficult challenge was to seamlessly integrate two heavily protected worlds: hospitals on one hand, and the grid, with minimal intrusiveness on both sides. This motivated a specialized architecture.

The architecture has two components: scheduler/worker agents at the user-level and the Interaction Bridge (IB) as an external service. The IB acts as a proxy between the PTM3D workstation, which is not gLite-enabled and the EGEE world. When opening an interactive session, the PTM3D workstation connects to the IB. In turn, the IB launches a scheduler and a set of workers on an EGEE site, through fully standard requests to an EGEE User Interface. A stream is established between the scheduler and the PTM3D front-end through the IB. When the actual volume reconstruction is required, the scheduler receives contours. The scheduler/worker agents follow a pull model with each worker computing one slice of the reconstructed volume at a time, and sending it back to the scheduler, which forwards them to IB from where they finally reach the front-end.

Results

The main result is that decoupling the local interaction and the grid computation gives an overall response time fully compatible with user requirements (less than 2 minutes), while the sequential time on a 3GHz PC with 2GB of memory can reach 20 minutes and more than 30 minutes on less powerful front-ends. The IB allows to tunnel a TCP flow; as a consequence, the local interaction scheme, which includes stopping, restarting and improving locally the segmentation, is kept strictly unmodified: when a contour is manually reshaped on the front-end, the new contours will be processed on the grid.

EVALUATING THE IMPACT OF COMPRESSION ON RIGID REGISTRATION ALGORITHMS

Lossy compression of medical images is a disputed point as the impact of compression algorithms on the use of images has to be carefully evaluated. In this section, we are interested in evaluating the impact of lossy compression on the performance of rigid registration algorithms (RAs). To realize this evaluation, we first determined the reference registrations from the uncompressed images using a statistical procedure called the *Bronze Standard*. Then, we evaluated to what extent the compression makes the registration results deviate from this reference.

The Bronze Standard Method

For the sake of analysis, one may consider registration algorithms as black boxes that take images as input and output a transformation. Considering the input data and the output transformation as random variables, the variability of the transformation (standard deviation or covariance matrix) can be propagated to some target points. Depending on the distance to the “true” transformation, one usually distinguishes gross errors (convergence to wrong local minima) and small errors around the exact transformation. The *robustness* quantifies the size of the basin of attraction of the right solution or the probability of false positives. The *accuracy* measures the error with respect to the truth (which may be unknown) for small errors. The *repeatability* or *precision* only measures the deviation from the average value, i.e. it does not take into account *systematic biases*, which are often hidden.

One of the simplest evaluation schemes is to simulate noisy data and to measure how far is the registration result from the true one (the ground truth is obviously known). The main drawback of synthetic data is that it is very difficult to identify and model faithfully all the sources of variability, and especially unexpected events (pathologies, artifacts, etc). Forgetting one single source of error (e.g. camera calibration errors in 2D-3D registration) automatically leads to underestimate the final transformation variability. Moreover, the faithful simulation of images is computationally extremely demanding. The second evaluation level is to use real data in a controlled environment, for instance imaging a physical phantom. However, one needs to measure precisely the motion or deformation of the phantom with an external apparatus to obtain a gold standard registration. Moreover, it is difficult to test all the clinical conditions (e.g. different types or localizations of pathologies). Last but not least, it is often argued that these phantoms are not representative of real in vivo biological system.

Thus, we are left with the last level of evaluation methods which relies on a database of in-vivo real images representative of the clinical application. Such a database should be large enough to span all sources of variability. In such a database, images are very realistic but establishing a gold standard registration is barely feasible. To still evaluate the performances of the registration algorithms without any gold standard, a first idea is to compose transformations that form a closed circuit and to measure the difference of the composition from the identity. This criterion does not require any ground truth, but it only measures the repeatability as any bias will get once again unnoticed.

In order to robustify the results, the idea of the Bronze Standard method is to build a statistical estimation of the ground truth by exploiting a large number of longitudinal image sequences and many different registration algorithms. The feasibility proof of the bronze standard technique was first established in (Roche 2001), but a proper deployment on a large database was only realized in (Glatard 2006) using the grid technologies presented in this paper. The principle is similar the STAPLE algorithm proposed for the validation of segmentation (Warfield 04): we consider the

exact registration as a hidden variable that we have to estimate concurrently with the variability of the registration results (e.g. the variance or mean square distance) with respect to these references. This is realized using an Expected Minimization like algorithm (Roche 2001).

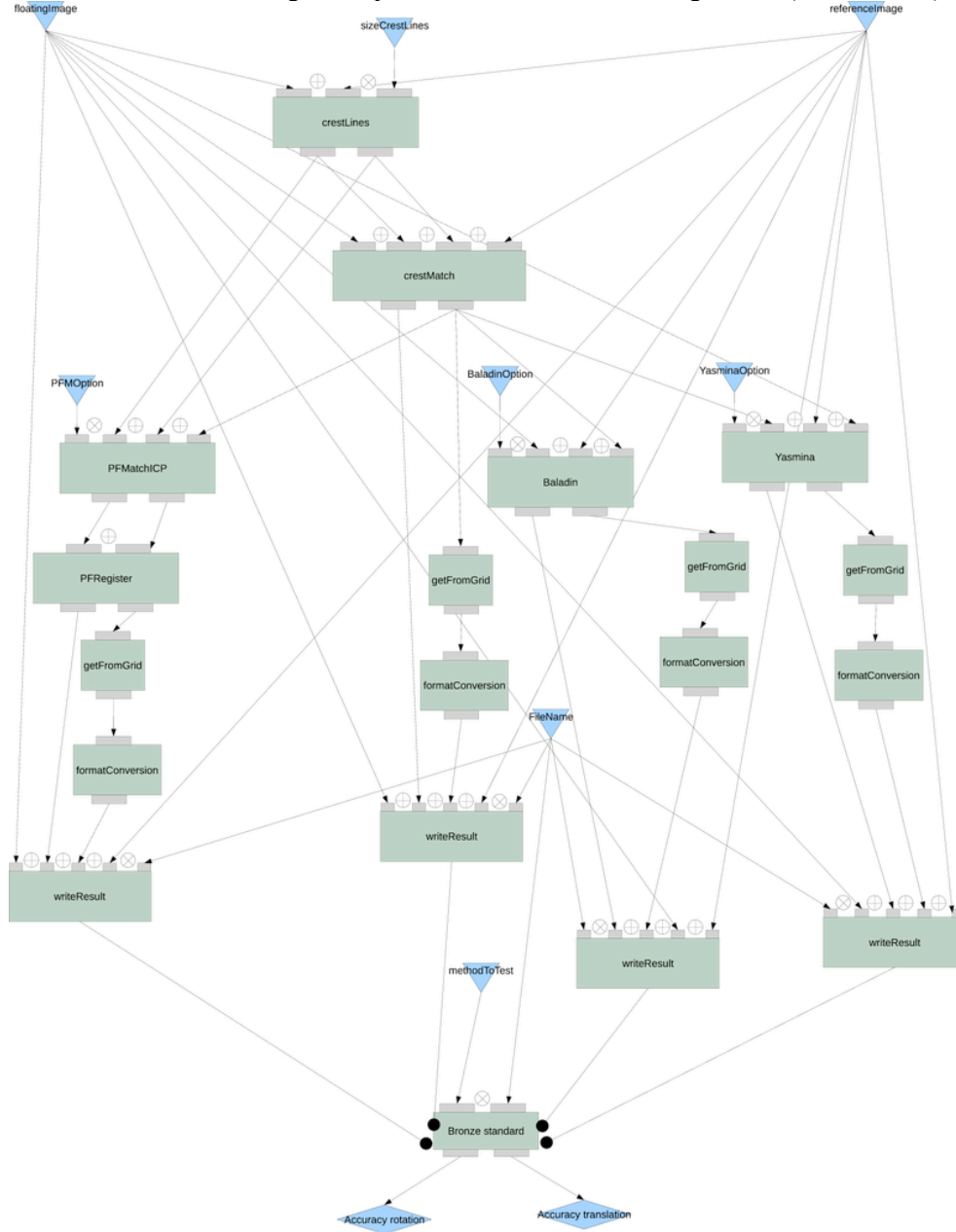


Figure 6. Bronze Standard workflow.

Application to the evaluation of the impact of the compression

The clinical problem context is the follow-up of brain radiotherapy. Monitoring tumor evolution requires a rigid registration of longitudinal images of the patient. The accuracy of this registration is directly related to the estimation of the tumor evolution. Experiments used a dedicated database of 65 images corresponding to 25 patients for which gadolinium injected MRIs have been acquired at several time points to monitor the tumors growth. A total of 126 registrations are required to relate all pairs images of the same patient. We used four different algorithms: *Baladin* (Ourselin 2000) and *Yasmina* (Roche, 2001) are intensity-based while

CrestLines and *PFRegister* (Pennec, 2000) are feature-based (extracting crest lines). The application workflow is displayed in figure 6.

This database has been compressed at Compression Ratio (CR) = 6,12,24,48 and 64, with the 3D-SPIHT algorithm. A total of 3024 transformations ($126 \cdot 4$ algorithms \cdot 6 CRs) have to be computed. Given the four registration algorithms, the total sequential execution time of this experiment is about 7.5 days. Through MOTEUR, the total duration of the experiment reduces to 18 hours on the shared EGEE production grid.

The number of outlier transformations gives an estimation of the robustness of the algorithms with respect to the compression. For each CR, outliers have to be identified by comparison to the transformations obtained from uncompressed images. Among the rejected transformations, a visual inspection is performed to determine whether they correspond to wrong local minima (when it is obvious that a manual registration can lead to a better result) or not. When a transformation is found to be in a wrong local minimum, the whole patient is removed. Otherwise, the absence of a specific algorithm for a given patient could bias the quantification of the accuracy of the remaining ones.

For each CR, the repeatability is measured by the variances of the transformations obtained from the compressed images only, after having removed the patients leading to a wrong local minimum in one of the CR.

The transformations obtained from the uncompressed images are considered as the reference for the evaluation. The accuracy of each algorithm is computed by measuring the mean distance of compressed transformations to the uncompressed reference. To avoid biases, the evaluated RA is excluded from the algorithms used to build the uncompressed reference.

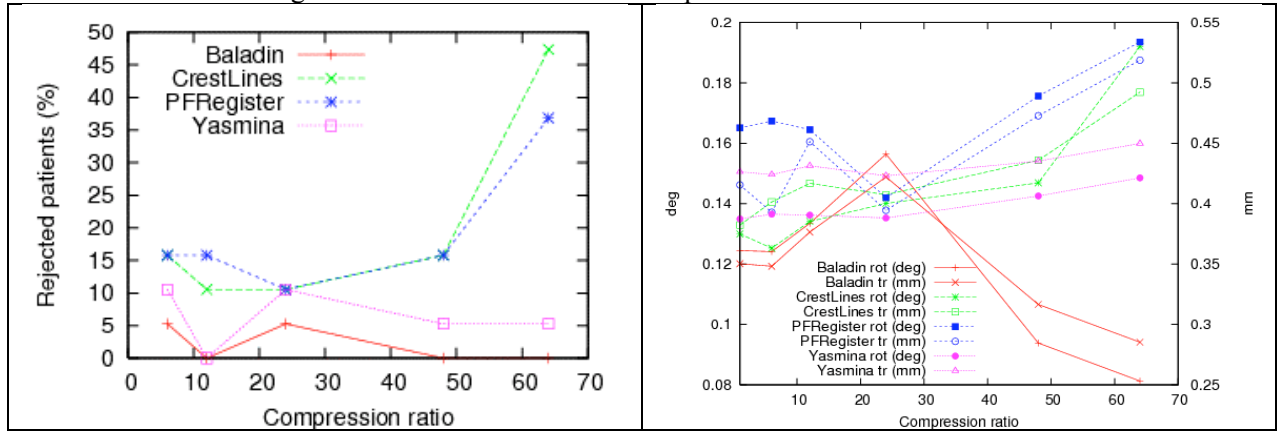


Figure 7. Left: Ratio of outlier patients with respect to the CR. One can see that the robustness does not significantly degrades until a CR of 64. Right: uncertainty of the algorithms measured by the standard deviation of the rotation (rot in degrees) and the translation (tr in millimeters) of each algorithm with respect to the uncompressed bronze standard.

Among the 25 patients of the database, 6 were removed as outliers. For these, the rigidity assumption was hardly valid, for instance because of high deformations in tumor areas. The reference was built from the uncompressed images of the remaining 19 patients. The ratio of outlier in those 19 patients is plotted on left of figure 7 for each algorithm. *Baladin* and *Yasmina* are very robust. It may be a consequence of their multi-scale strategy (both use a pyramidal decomposition). For those two algorithms, it seems that the robustness even improves with the CR. For *CrestLines* and *PFRegister*, the robustness degrades significantly for a CR higher of 64.

This was expected since the image derivatives used to extract features are highly impacted by the compression.

Figure 7 right displays the uncertainty of the algorithms measured by the standard deviation of the rotation and the translation components of the transformation. As it measures the distance to the uncompressed case, the smaller the better. The uncertainty of feature-based methods is highly augmented at CR=64. *Yasmina* is quite insensitive to the compression: its mean error only increases by 10% for the rotation and by 5.5% for the translation. More surprisingly, after a brief rise until CR 24, the accuracy of *Baladin* is improving: at CR=64, it is 34% better than without compression for the rotation and 18.5% better for the translation. This behavior may be due to the multi-scale handling compensating the effects of potential noise introduced in the images.

In summary, results tend to show that the impact of 3D-SPIHT compression on robustness, repeatability and accuracy is quite negligible until a significant CR (48). Beyond this threshold, feature-based methods based on crest-lines are highly penalized, while intensity based algorithm do not seem to be deeply impacted, probably thanks to a good multi-scale handling. Results on *Baladin* are interesting as they tend to show that performances can even improve.

Drawing general conclusions about the impact of lossy compression methods on registration algorithm would still require to confirm these trends with other registration problems and compression methods. Lossy compression in medical applications should be evaluated by experienced radiologists over a large data set. The fact that this kind of study is costly to set up, and time-consuming for the experts, explains the scarcity of the literature on the actual 3D lossy coders that uses a medical opinion. With the application of the Bronze Standard method to evaluation of lossy compression, we propose an intermediate step, which automatically builds an approximation of the ground truth, and could be an incentive towards targeted studies involving clinicians.

CARDIAC IMAGE ANALYSIS

Motivation

Cardiac imaging has gained a central place in the diagnosis and treatment of cardiac pathologies, which are still the leading cause of death in industrial countries. Examinations in MRI or X-Ray CT typically result in a huge amount of 3D/4D image data delivering kind of snapshots of the heart status at successive time points of the cardiac cycle. In MRI, a set of 3 to 10 parallel slices are usually acquired over the heart. At each slice level, a dynamic series of frames is acquired during the cardiac cycle. This acquisition represents about 400 MBytes of 3D interpolated image data. A complete examination on a patient usually requires tens of them. The extraction of quantitative functional parameters, needed for cardio-vascular disease diagnosis, relies on sophisticated analysis programs that may require intensive computational resources in order to produce results in time compatible with the clinical practice. In particular, a method has been designed to retrieve the heart's motion from 3D dynamic MR imaging data. Indeed, motion appears to be one of the altered functions in severe cardiac ischemic diseases.

The motion estimation problem is a typical example of inverse problems which necessitates high level image processing methods with the introduction of *a priori* models and information. We have defined a method where the estimation is performed within a state space formulation and a recursive Kalman filtering approach (Delhay, 2006; Delhay, 2007). It takes into account the whole image sequence achieving a spatiotemporal (ST) modeling of the motion flow over the cardiac cycle. Moreover, it can be rather straightforwardly parallelized. With the aim of routinely

processing cardiac examinations in a clinical setting, one needs high performance computing facilities such as those offered by computational grids.

The ST model principle

The principle of our motion estimation approach is closely related to the optical flow methods and their extension to multiple frames. The spatiotemporal motion is modeled by a parametric free form deformation (FFD) based model in both 3D space and time (Sederberg, 1986). A regular grid of control points (CP) encloses the initial space at rest. L is the number of grid nodes. The displacement of control points implies a dense spatial deformation the effect of which depends on the FFD function basis.

We proposed to extend the initial spatial approach to a spatio-temporal approach by attributing to each CP a continuous trajectory. The collection of all the CP trajectories forms the ST model parameters; each trajectory is individually modeled by a Fourier series of order N . This formulation insures two main properties of our model: the continuity and periodicity of the motion over the cardiac cycle. Therefore, the number of model parameters is $3xLx(2N+1)$ in 3D space. At the finest transformation level the grid size is currently of size $19x19x19$ and the ST model is completely defined by 185193 parameters with $N=4$, which makes the problem quite difficult to solve. Thanks to the periodization of the image sequence, the estimation is performed within a state space formulation and a recursive Kalman filtering approach (Delhay, 2006; Delhay, 2007). The goodness of fit is estimated through the sum of squared difference similarity criterion

Parallelization strategies.

Due to the high complexity of the problem, especially in the 3D+time space, several strategies have been proposed and implemented to reduce the computing time. First, a double multi-level approach has been adopted: a multi-resolution representation of the images (considering images with increasing details) and a multi-level representation for the transformations. Also, strategies have been set up to focus on moving regions and automatically tune the temporal model order through the computation of attention maps (Delhay, 2007). The process allows for a natural splitting for parallel computations without impacting the result accuracy. As shown in figure 8, the computational load can be distributed on multiple processors that all receive part of the image domain. Processes are synchronized to reconstruct the whole result. The partitioning is performed so as to balance the computational load of each processor in order to reduce latency as much as possible. Multi-threaded and message-passing (MPI) versions have been developed and deployed on a shared memory multiprocessors (SMP) machines and on PC clusters, resulting in an almost linear relation between computing time and CPU number (figure 8).

In the view of going one step further in the context of multi-center clinical studies on large patient populations, the method has also been ported to the EGEE grid. The grid version exploits trivial parallelism (across patients) as well as coupled parallelism (MPI on PC clusters for each patient dataset).

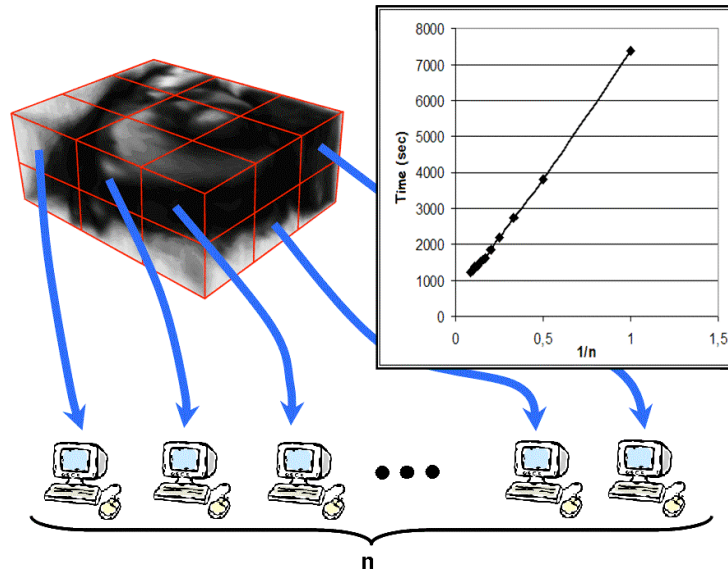


Figure 8. Splitting computation load among processors obtained by partitioning the image data and an example of the computing time as a function of the processor number (here on a SMP architecture).

Results and discussion

The experiments conducted both on synthetic and true MR image sequences demonstrated the good accuracy of the proposed method. On a 2D+time synthetic realistic sequence for instance, the accuracy (over the whole sequence) is $1,8 \pm 1,2$ mm with $L = 21$ and $N = 5$. However, model identification is clearly a time-consuming process. A 4D experiment required 20 CPU hours for 160^3 image size and 21 frames, on a single processor machine. Experiments on a SMP-SGI system confirmed that the computation time reduces almost linearly as a function of the number of processors (approximately 2 hours in the previous case with 10 processors). Currently, the MPI version is evaluated on the EGEE grid. Our intention is to make the method available to the community through a web portal and evaluate its performance on a clinical database before deploying it for patient specific analysis and large multi-center studies on cardiac pathologies. Also, the complete sequence of computations, including the various image processing operators involved, has been reformulated within workflow representation languages and partly implemented on two workflow engines, including MOTEUR, in the context of the French ANR project GWENDIA (<http://gwendia.polytech.unice.fr>).

HUMANITARIAN MEDICINE

One of the greatest opportunities offered by grids is to build federation of databases where the medical data are kept locally in the different participating hospitals and can be queried remotely on request by clinicians for second diagnosis. This section describes a platform developed for this purpose on top of the gLite data management services using Web Service technology.

Introduction to grid-enabled telemedicine

The concept of telemedicine covers all the potential means to exchange medical data or images whatever the distance between two physicians,. As stressed in a very recent report from the European Commission and the World Health Organization, “Information and communication technologies are changing health care delivery and are at the core of effective, responsive health systems. These technologies are key to connecting people, information and research to improve

health in countries. They are also vital in enabling rapid response to global threats to health.” (Dzenowagis, 2005)

One of the main added values of grids for telemedicine is to offer to healthcare professionals the capacity to query distant databases in a transparent and secure way (Gonzales, 2005).

- Patient data are kept in databases inside hospitals where patients are treated.
- Access for second diagnosis can be strictly controlled through the authorization and authentication mechanisms of the grid. Only a subset of the patient record can be exposed through the grid.
- In case additional information is needed, the team of clinicians requesting the second diagnosis can provide access to more patient information or perform new examinations and store the results on their local database.
- The data being distributed over the participating hospitals, there is no single point of failure

Platform architecture and design

The platform goal is to enable the exchange and usage of medical information over the grid in a secure way, which means exposing the information to the grid users which are allowed to access it and exploiting the medical information with analysis and job management facilities.

The platform was designed to fulfill additional requirements:

- The applications running on the platform must be easy to use as healthcare professionals are too busy to learn complex tools. In particular, the grid complexity must be completely hidden.
- The platform should be adaptable to any medical field. It must be flexible so it can be customized to adapt to the “modus operandi” of its users.
- The platform must be able to interact with existing hospital data and image archiving systems (PACS).
- The platform should be fault tolerant in the perspective of clinical routine
- The platform should not require global connectivity. Any network failure outside the hospital should not affect data management in the hospital, nor the exchange and treatment of the medical information outside the area affected by the failure.
- The platform should fulfill all legal requirements in relation to security

The platform architecture is built on three pillars:

- The medical data are stored in an AMGA server inside the hospital. AMGA is the metadata catalogue developed within the EGEE project (see section)
- The healthcare professional accesses all the platform services through a web portal developed with the GridSphere portlet container
- The grid infrastructure provides the storage and computing resources needed to store the medical images and to treat the data according to the user need.

Figure 9 displays the overall architecture of the platform (left vignette) and shows how information is shared between hospitals in different locations (right vignette).

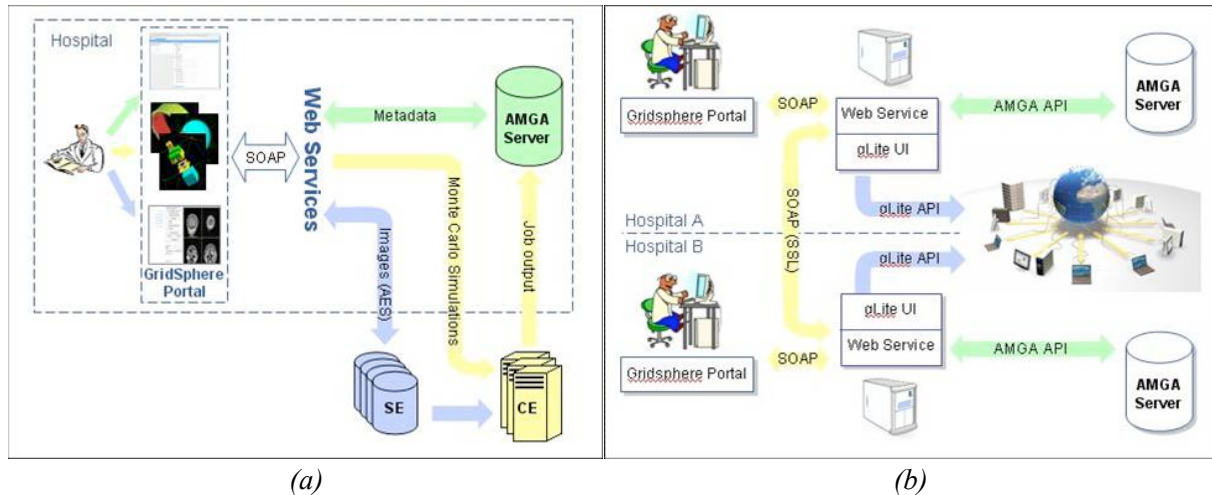


Figure 9. (a) Information and image management inside a hospital; (b) information and image sharing between hospitals in different locations.

In the design of the telemedicine, a separation was made between the presentation layer, developed using Gridsphere portal container, and the business logic part which was developed using web services to manage all the low level interaction with different grid services. This also allows the design of a personalized platform interface for each different customer without changing anything on data management or grid interaction side.

The Medical Data Manager described in section 3 was not deployed on the EGEE Infrastructure when the development of the platform was started. As a consequence, only one its component, namely the AMGA metadata catalogue, was used. However, the platform design allows to easily interface it to the MDM services.

The choice of web service technology came from the need to use standard and well supported communication protocols (SOAP in our case) to guarantee interoperability and security between different instances of the platform installed in remote hospitals. The web service layer manages all routing mechanisms used to exchange SOAP messages between remote platform instances allowing physicians belonging to different and remote hospitals to share patient files, clinical exams and medical images.

Medical information and DICOM metadata extracted from medical images are managed using AMGA metadata catalogue. This grid service provides access to different database back-ends in a uniform way and independently of the particular database system used. This makes AMGA a really interesting service to develop software solutions in a distributed and heterogeneous environment.

To adapt to the existing hospital information systems, the architecture should allow different hospitals to specify a different (and usually personalized) database structure for the same medical field. To make this possible we decided to use the XML mark-up language to represent all the information managed and exchanged in the platform. Together with SOAP protocol and the Web Service technology, the use of XML documents to store and exchange information makes the platform completely independent and decoupled from any particular database structure deployed inside each hospital.

Platform deployment

The platform beta version is currently deployed in four sites; three of them are research laboratories while one is a hospital in Clermont-Ferrand. The platform has been customized for

an application in the field of neurosurgery. Indeed, since March 2000, humanitarian missions have been conducted in neurosurgical unit of the Regional Hospital of District of Chuxiong (Yunnan, PRC) by French clinicians from healthcare centers in Clermont-Ferrand in order to operate neurosurgical pathologies including hydrocephalus in children which were not treated previously in this hospital. As these missions are limited in time, key issues for their success are the identification of patients needing surgery before the mission and their post-surgery follow-up after the mission. Another issue is related to the transfer of skills and techniques which should not be limited to the mission durations.

The HOPE platform was customized to the needs of the clinicians and designed in collaboration with them. A standard medical file was designed and translated in Chinese language. This file, subdivided in six sheets, gathers the relevant information at the different steps of the medical procedure (parental agreement for surgery, pre-operative status, operative report, follow-up, complications and re-operation, anesthetic examination). However, due to network problems, its deployment has not yet been possible in Chuxiong.

CONCLUSION AND PERSPECTIVES

Parallel computer architectures and algorithms have co-evolved with hard science applications, on an ascending ladder of trials and improvements. At the opposite, researchers in medical imaging know all too well that imperfect but promising, or simply new and not fully validated methods and algorithms, have little hope to go from scholarly publication into clinical research. Considering the opportunities offered by modern computing infrastructures, clinical research and practice is not fully aware that high performance computing may be more than the limited parallelism offered by a high-end workstation.

Amongst the many causes of this sobering fact, the first one is sociological: given the academic structures, the opportunities to gather truly multi-disciplinary teams in medical image analysis like e.g. the BIRN project are much less frequent than in the hard science field. At the French level, AGIR has gathered a truly multi-disciplinary team, involving computer scientists with various specialties, and physicians. As a consequence, one of the basic options of AGIR is that the medical requirements regarding computing practice (dependability, interactivity, security and privacy protection, compliance to the DICOM standard) must be considered as first-order requirements, not as user interface or practical issues. In the limits of a 3-years project, AGIR has followed the path pioneered at the Surgical Planning Laboratory (Kikinis, 1998): “utilization of HPC technology (...) in an application driven clinical and research environment”. Dependability requirements oriented AGIR to focus the core medical services towards production grids, and namely the EGEE grid, which design was largely driven by a very different community. The responsiveness and data management have successfully implemented the basic services complying with these requirements on top of the EGEE grid.

A second cause is that the medical constraints have very limited flexibility: a physicist or a chemist can appreciate the potential of prototype tools, and experiment on it, providing feedback to the computer scientist; a physician requires highly dependable, fully secure, user-friendly software, which has also to comply with complex standards. Software dependability involves both production quality at the usage level, and guarantees about the medical relevance of the embedded analysis tools, which is mainly about validation. Production grids change the conditions of validation of methods and algorithms. Validating new processing methods requires large computing power, large datasets, or both. With production grids, the infrastructure for sharing images and algorithms is here, together with the availability of computing power increased by orders of magnitude with respect to the usual workstation. At the same time, the

development and generalization of service-oriented frameworks offers the perspective of integrating the grid resource as a backend of the usual interfaces and workflows of the physician. Ideally, the computer scientist will focus on its core contribution, methods and algorithms, and the physician could evaluate them into existing and appreciated environments. The application of the Bronze standard method to compression evaluation is an excellent example of integration of these multiple expertise, whose enactment is made possible by the contributions of the core medical services, and finally the grid infrastructure.

Overall, the main results of the AGIR project are twofold: real-world production oriented grids, and namely the EGEE grid, allow breakthrough in medical image processing research; second, some of the services required for clinical research and practice can be deployed on production grids.

The interaction with the medical partners has been pervasive all along the project, by providing real (and thus difficult) datasets, and has continuously steered the project towards the clinical constraints. To some extent, the value of their contribution is demonstrated by the fact that the experience and tools developed in AGIR has founded the participation of AGIR partners in two much larger-scale projects in the area of grid-enabling medical image processing: the FP6 IP Health-e-child, dedicated to clinical research in pediatric healthcare, and the French software R&D project NeuroLOG, dedicated to open-source grid-enabled tools for knowledge integration, with applications to neuro-imaging and oncology.

ACKNOWLEDGEMENTS

The AGIR project has been funded by the French Ministry of Research under the *Masses de Données* program. Developments has been partially supported by the EGEE-II project funded by the European Union INFSO-RI-031688, by other grants from the European Commission (BioinfoGRID, Embrace), and for the work in section 7 by regional authorities (Conseil Régional d'Auvergne, Conseil Général du Puy-de-Dôme, Conseil Général de l'Allier). The authors are indebted to M. Diarena, B. Grenier, A. Fessy, S. Nowak, N. Verhaeghe and N. Spalinger for their invaluable contribution to the development of the HOPE platform and its customization for telemedicine.

REFERENCES

Allcock, B., Bester, J., Bresnahan, J., Chervenak, A.L., Foster, Ian, Kesselman, Carl, Meder, S., Nefedova, V., Quesnal, D. & Tuecke, S. (2002). Data Management, Transfer in High Performance Computational Grid Environments. *Parallel Computing Journal*, 28 (5). 749-771.

Aloiso, G., Benkner, S., Bilofsky, H., Blanquer Espert, I., Breton, V., Cannataro, M., Chouvarda, I., Claerhout, B., Dean, K., Fiore, S., Hassan, K., Heeren, G., Hernández García, V., Herveg, J., Hofmann, M., Jones, C., Koutkias, V., Lloyd, S., Lonsdale, G., López, V., Maglaveras, N., Maigne, L., Malousi, A., Martin-Sanchez, F., McLatchey, R., Medico, E., Miguët, S., Mirto, M., Montagnat, J., De Moor, G., Nozaki, K., De Neve, W., Oliviera, I., Pennec, X., Sanchez, J., Solomonides, T., Taillet, M., Veltri, P., De Wagster, C. & Ziegler, R. (2005). HealthGrid White Paper, <http://whitepaper.healthgrid.org>.

Atkins, D. (2003). Report of Blue-Ribbon Advisory Panel on Cyberinfrastructure. <http://www.nsf.gov/publications>.

Baruah, S., Cohen, N., Plaxton, C. & Varvel, D. (1996). Proportionate Progress: A Notion of Fairness in Resource Allocation. *Algorithmica*, 15(6), 600-625.

- Basu, S., Talwar, V., Agarwalla, B., & Kuma, R. (2003). *Interactive Grid Architecture for Application Service Providers*. HP Technical Report HPL-2003-84R1.
- Berman, F. Wolski, R. & Casanova, H. (2003). Adaptive Computing on the Grid Using AppLeS. *IEEE Transactions on Parallel and Distributed Systems*, 14(4), 369–382.
- Berry, D., Germain-Renaud, C., Hill, D., Pieper, F. & Saltz, J. (2004). Report on the Workshop IMAGE'03: Images, medical analysis and grid environments. TR UKeS-2004-02, UK National e-Science Centre.
- Bilgin, A., Zweig, G., & Marcellin, M.W. (2000). Three-dimensional image compression with integer wavelet transforms. *Applied Optics*, 39, 1799-1814.
- Bolze, R. Cappello, F., Caron, E., Daydé, R., Desprez, F., Jeannot, E., Jégou, Y., Lanteri, S., Leduc, J., Melab, N., Mornet, G., Namyst, R., Primet, P., Quetier, B., Richard, O., Talbi, EG. & Touché I. (2006) Grid'5000: a large scale and highly reconfigurable experimental Grid testbed. *International Journal of High Performance Computing Applications*, 20(4):481-494.
- Brevik, J., Nurmi, D. & Wolski, R. (2006, March). *Predicting bounds on queuing delay for batch-scheduled parallel machines* (pp 110–118). 11th ACM SIGPLAN symposium on principles and practice of parallel programming. New York, NY: ACM Press.
- Caron, E., & Desprez, F. (2006) DIET: A Scalable Toolbox to Build Network Enabled Servers on the Grid. *International Journal of High Performance Computing Applications*, 20(3), 335-352.
- Cho, S., Kim, D., & Pearlman, W.A. (2004). Lossless compression of volumetric medical images with improved three-dimensional SPIHT algorithm. *Journal of Digital Imaging*, 17, 57-63.
- Delhay, B., Clarysse, P., Pera, C. & Magnin, IE. (2006). A spatio-temporal deformation model for dense motion estimation in periodic cardiac image sequence. *From Statistical Atlases to Personalized Models: Understanding Complex Diseases in Populations and Individuals, Satellite Workshop MICCAI 2006*. Copenhagen, Denmark.
- Delhay, B., Clarysse, P., Pera, C. & Magnin, IE. (2007, June). *Locally adapted spatio-temporal deformation model for dense motion estimation in periodic cardiac image sequences*. Functional Imaging and Modeling of the Heart (pp. 393-402). Salt Lake City, UT: Springer-Verlag Lectures Notes on Computer Science 4466.
- Dinda, P. (2006). Design, Implementation, and Performance of an Extensible Toolkit for Resource Prediction In Distributed Systems. *IEEE Transactions on Parallel and Distributed Systems*, 17(2), 160-173.
- Dzenowagis, J. & Kernen, G. (2005). *Connecting for Health: global vision, local insight*. Report for the World Summit on Information Society. Tunis, Tunisia: World Health Organization Press
- EGEE SDJ WG .(2006). Report of the Short Deadline Jobs Working Group. <http://egee-na4.ct.infn.it/wiki/index.php/ShortJobs>.
- Erberich, S., Silverstein, J., Chervenak, A., Schuler, R., Nelson, M. & Kesselman, K. (2007a). Globus MEDICUS – Federation of DICOM Medical Imaging Devices into Healthcare Grids. *Studies in Health Technology and Informatics*, 126, 269-278.

Erberich, S., Silverstein, J., Kratz M., Chervenak, A., Schuler, R., Nelson, M. & Kesselman, K. (2007b, April). *Globus MEDICUS – Standards Based Enterprise for Medical Image Publication, Discovery, and Archiving in HealthGrids*. Paper presented at the Open Source Solutions for Multi-Center Information Management. Saint Louis, USA.

Erickson, B. (2002). Irreversible Compression of Medical Images. *Journal of Digital imaging*, 15(1), 5-14.

Foster, I. & Kesselman, K. (Eds) (2003). *The Grid 2: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers.

Foster, I., Kesselman, K. & Tuecke, S. (2001). The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *International Journal Supercomputer Applications*, 15(3), 200-222.

Foster, I. (2006, October) Globus Toolkit Version 4: Software for Service-Oriented Systems. *IFIP International Conference on Network and Parallel Computing* (pp 2-13). Tokyo, Japan: Springer-Verlag Lectures Notes on Computer Science 3779.

Gagliardi, F., Jones, B., Grey, F., Bégin, ME. & Heikkurinen, M. (2005). Building an Infrastructure for scientific Grid computing: status and goals of the EGEE project. *Philosophical Transactions of the Royal Society A*, 363(1833), 1729-1742.

Gaudeau, Y., & Moureaux, J.M. (2007, November). *Lossy compression of volumetric medical images with 3D dead zone lattice vector quantization*. 26th Picture Coding Symposium. Lisboa, Portugal.

Germain-Renaud, C., Perez, J., Kégl, B. & Loomis, C. (2008, May). Grid Differentiated Services: a Reinforcement Learning Approach. *8th IEEE Symposium on Cluster Computing and the Grid*. Lyon, France: IEEE Computer Society.

Germain-Renaud, C., Loomis, C., Moscicki, JT., & Texier. R. (2008). Scheduling for responsive grids. *Journal of Grid Computing*, 6(1), 15-27.

Glatard, T., Pennec, X. & Montagnat, J. (2006). Performance evaluation of grid-enabled registration algorithms using bronze-standards. In *Proc. of MICCAI'06, Part II*, LNCS 4191, pages 152-160.

Glatard, T., Montagnat, J., Erberich S., Carmeli B. & Kaufman, J. (2007, November). *Cross Enterprise Document and Image Exchange (XDS/-I) in HealthGrids*. Paper presented at the Radiology Society of North America (RSNA'07). Chicago, USA.

Glatard, T., Montagnat, J., Emsellem, D. & Lingrand, D. (2008a). A Service-Oriented Architecture enabling dynamic services grouping for optimizing distributed workflows execution. To appear in *Future Generation Computer Systems*.

Glatard, T., Montagnat, J., Lingrand, D. & Pennec, X. (2008b). Flexible and efficient workflow deployment of data-intensive applications on grids with MOTEUR. To appear in *International Journal of High Performance Computing and Applications*.

Gonzales, J., Pomel, S., Breton, V., Clot, B., Gutknecht, J.L., Irrthum, B. & Legré, Y. (2005) Empowering humanitarian medical development using grid technology. *Methods of Information in Medicine*, 44, 186-189.

Heinzlreiter, P. & Kranzlmueeller, D. (2003). Visualization Services on the Grid: The Grid Visualization Kernel. *Parallel Processing Letters*, 13(2), 135–148.

Jeannot, E. (2005, April), Improving middleware performance with AdOC: an Adaptive Online Compression library for data transfer. *IEEE International Parallel and Distributed Processing Symposium* (pp 70). Denver, CO, USA: IEEE Computer Society.

Jensen, ED., Locke, CD. & Tokuda, H. (1985, December). *A time driven scheduling model for real-time operating systems* (pp 112–122). IEEE Real-Time Systems Symposium, San Diego, CA : IEEE Computer Society.

Keahey, K., Foster, I., Freeman, T. & Zhang, X. (2006). Virtual Workspaces: Achieving Quality of Service and Quality of Life in the Grid. *Scientific Programming*, 13(4), 265-275.

Kikinis, R. Warfield, S. & Westin, C. (1998). *High Performance Computing (HPC) in Medical Image Analysis (MIA) at the Surgical Planning Laboratory (SPL)*. Paper presented at the 3rd High Performance Computing Asia Conference & Exhibition, Singapore, Malaysia.

Laure, E. et al. (2004, September). *Middleware for the next generation Grid infrastructure*. Paper presented at the 2004 Conference for Computing in High Energy and Nuclear Physics, Interlaken, Switzerland.

Leenaars, M., Heikkurinen, M., Louridas, M., & Karayannis, F. (2005). e-infrastructures roadmap. Published by e-IRGSG. <http://www.e-irg.org/roadmap/eIRG-roadmap.pdf>

Lin,, B. & Dinda, PA. (2005, November). *VSched: Mixing Batch And Interactive Virtual Machines Using Periodic Real-time Scheduling*. ACM/IEEE conference on Supercomputing: IEEE Computer Society

Montagnat, J., Bellet, F., Benoit-Cattin, H., Breton, V., Brunie, L., Duque, H., Legré Y., Magnin, I., Maigne, L., Miguet, S., Pierson, J.-M., Seitz, L. & Tweed, T. (2004). Medical images simulation, storage, and processing on the european datagrid testbed. *Journal of Grid Computing*, 2 (4), 387-400.

Montagnat, J., Frohner, Á., Jouvenot, D., Pera, C., Kunszt, P., Koblitz, B., Santos, N., Loomis, C., Texier, R., Lingrand, D., Guio, P., Brito Da Rocha, R., Sobreira de Almeida, A. & Farkas, Z. (2008). A Secure Grid Medical Data Manager Interfaced to the gLite Middleware. *Journal of Grid Computing*, 6 (1), 45-59.

Moore, GA. (1991). *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*. Harper Business.

Oinn, T., Addis, M., Ferris, J., Marvin, D., Senger, M., Greenwood, M., and Carver, T., Glover, K., Pocock, M. R., Wipat, .A & Li, P. (2004). Taverna: A tool for the composition and enactment of bioinformatics workflows. *Bioinformatics journal*, 17 (20), 3045-3054.

Osorio A., Valette PJ., Mihalcea A., Atif J. & Ripoche X. (2002) *A new PC based software to perform semi-automatic hepatic segmentation using CT and MR images*. Presented at InfoRAD-RSNA, Chicago, USA.

Ourselin, S., Roche, A., Prima, S. & Ayache, N. (2000). Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images. In proc. of *MICCAI 2000*, Lectures Notes in Computer Science 1935, Pittsburgh, Penn, USA, pages 557-566. Springer.

Peltier, ST. & Ellisman MH. (2003) The Biomedical Informatics Research Network (BIRN). In: I Foster and C Kesselman (Eds), *The Grid: Blueprint for a New Computing Infrastructure* (pp 109-120). Morgan Kaufmann Publishers.

Penedo, M., Souto, M., Tahoces, P.G., Carreira, J.M., Villalón, J., Porto, G., Seoane, C., Vidal, J.J., Berbaum, K.S., Chakraborty, D.P., & Fajardo, L.L. (2005). Free-response receiver operating characteristic evaluation of lossy JPEG2000 and object-based set partitioning in hierarchical trees compression of digitized mammograms. *Radiology*, 237, 450-457.

Pennec, X & Thirion, J-P. (1997). A Framework for Uncertainty and Validation of 3D Registration Methods based on Points and Frames. *Int. Journal of Computer Vision*, 25(3):203-229.

Pennec, X., Ayache, N. & Thirion, J.P. (2000). Landmark-based registration using features identified through differential geometry. Chapter 31 of *I. Bankman, editor, Handbook of Medical Imaging*, , pages 499-513. Academic Press.

Perelmotov, T., Petravick, D., Gu, J., Barring, O., Baud, J.-P., De Witt, S., Jensen, J., Synge, O., Haddox-Schatz, M., Hess, B., Kowalski, A. and Watson, C. (2002). SRM Interface Specification v2.2. *FNAL technical report*. May 2002. <http://sdm.lbl.gov/srm-wg/doc/SRM.v2.2.html>

Pujol, S, Kikinis, R & Gollub, R. (2008). Lowering the barriers inherent in translating advances in neuroimage analysis to clinical research applications. *Acad Radiology*, 15(1), 114-118.

Roche A., Pennec X., Malandain G. & Ayache N. (2001). Rigid Registration of 3D Ultrasound with MR Images: a New Approach Combining Intensity and Gradient Information. *IEEE Transactions on Medical Imaging*, 20(10):1038-1049.

Santos, N. & Koblitz, B. (2006, June). *Distributed Metadata with the AMGA Metadata Catalog*. Paper presented at the workshop on Next-Generation Distributed Data Management, Paris, France. <http://project-arda-dev.web.cern.ch/project-arda-dev/metadata>

Sederberg, T. W., & Parry, S. R. (1986). Free-form deformation of solid geometric models. *ACM Computer Graphic*, 20, 151-160.

Shamir, A. (1979). How to share a secret. *Communications of the ACM*, 22, 612-613.

Sim, A. & Shoshani, A. (2007) *The Storage Resource Manager Interface Specification Version 2.2* <https://forge.gridforum.org/sf/projects/gsm-wg>

Snell, Q., Clement, MJ., Jackson DB. & Gregory, C. (2000, May). *The Performance Impact of Advance Reservation Meta- scheduling*. 6th. workshop on Job Scheduling Strategies for Parallel Processing (pp 137–153), Cancun, Mexico: Springer-Verlag Lecture Notes on Computer Science 1911.

Tesauro, GJ. & Kephart, JO. (2004, May). *Utility functions in autonomic systems*. 1st International Conference on Autonomic Computing (pp. 70–77). New York, NY: IEEE Computer Society.

Texier, R (2007) *Systèmes de globalisation des ressources données-calculs. Application à l'imagerie médicale*. Unpublished doctoral dissertation, Université Paris-Sud, France.

Traxer, O., Osorio A., Pasqui, F., Sebe, P., Merran, S., Ripoche, X., Gategno, B. & Thibault, P. (2005, March). *New PC based software using an Augmented Reality system to simulate PCNL*. 20th European Association of Urology (EAU) Congress. Istanbul, Turkey.

Tsaregorodtsev, A., Garonne, V. & Stokes-Rees, I. (2004). DIRAC: A Scalable Lightweight Architecture for High Throughput Computing (pp 19-25). 5th IEEE/ACM International Workshop on Grid Computing: IEEE Computer Society.

Warfield S.K., Zou, K.H, & Wells, W.M. (2004). Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation. *IEEE Transactions on Medical Imaging*, 23(7) 903-921

Xiong, Z., Wu, X., Cheng, S., & Hau, J. (2003). Lossy-to-Lossless compression of medical volumetric data using three-dimensional integer wavelet transform. *IEEE Transactions on Medical Imaging*, 22, 459-470.

KEY TERMS & DEFINITIONS

Medical imaging: the techniques and processes used to create digital images of the human body (or parts of it) for clinical purposes (diagnosis and therapy) or medical science (including the study of normal anatomy and function).

Performance evaluation: assessment of the technical characteristics of algorithms used in medical image analysis (e.g. registration accuracy). In some cases, a well established reference or gold standard exists and can be used for evaluation. In other cases, there is no direct way of comparing a result to the known reference. Statistical procedures can be used to define an approximate reference, or Bronze Standard.

Image registration: evaluating the transformation compensating for the difference in location, pose (rigid registration) and possibly scale or shape (non-rigid registration) between the images. Different images can thus be spatially aligned in one common coordinate system before further analysis. For instance, rigid registration of images of the head is used to correct for the different positioning of the subject in the image acquisition device at two distinct times points.

Data compression : encoding information using fewer bits than the unencoded representation of data would. Lossy compression methods do not allow to fully retrieve the original data, while lossless compression methods do.

DICOM (Digital Imaging and Communications in Medicine): DICOM is standards to ensure the interoperability in the processing of medical image. In particular, DICOM defines standard for image and data formats, image transmission, query or store images, print and display.

Telemedicine: clinical practice where medical information is transferred via telephone, the Internet or other networks for the purpose of consulting, and sometimes remote medical procedures or examinations.

Responsiveness: in the context of human-computer-interaction, responsiveness of an interactive system describes how quickly it responds to user input.

Workflow: a description of a process involving the ordered execution of different tasks. A workflow graph describes the tasks involved and their

dependencies. The dependencies might be temporal or due to data exchanges needed between tasks (data flow).

Quality of Service: the ability to provide different priority to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow.